

Ідентифікація Людини за її ЕКГ з Допомогою Зростаючої Пірамідальної Мережі

Лілія Кізуб

Інститут проблем математичних машин і систем
НАН України
Київ, Україна
lilimmsp@gmail.com

Віталій Величко

Інститут кібернетики ім. В.М. Глушкова
НАН України
Київ, Україна
velychko@aduis.com

Human Identification Based on Electrocardiogram Using Growing Pyramidal Network

Liliia Kizub

Institute of Mathematical Machines and
Systems of NASU
Kyiv, Ukraine
lilimmsp@gmail.com

Vitalii Velychko

V.M. Glushkov Institute of
Cybernetics of NASU
Kyiv, Ukraine
velychko@aduis.com

Анотація—У статті розглядається вирішення задачі ідентифікації людини за її електрокардіограмою, з допомогою класифікатора, що базується на використанні зростаючих пірамідальних мереж. При наявності відповідних ознак електрокардіограми (ЕКГ) – координат точок канонічного сплайну, що апроксимує тривимірну криву QRS - комплексу, може бути проведена класифікація ЕКГ. Представлені результати проведених експериментів.

Abstract—The paper deals with the solution of the problem of identifying a human by its electrocardiogram, using a classifier based on the use of growing pyramidal networks. In the presence of the corresponding features of the electrocardiogram (ECG) - the coordinates of the points of the canonical spline, which approximates the three-dimensional curve of the QRS complex, an ECG classification can be performed. The results of conducted experiments are presented.

Ключові слова—електрокардіограма (ЕКГ); зростаюча пірамідальна мережа; QRS-комплекс, характерні ознаки ЕКГ, ідентифікація, автентифікація

Keywords—electrocardiogram (ECG); growing pyramidal network; human identification; QRS-complex; characteristic

Робота підтримана грантом НАН України «Розвиток сервісів ВО «МедГрид» для підтримки мобільних діагностичних платформ та персоналізованої діагностики на базі електрокардіографії та спірометрії. Другий рік виконання» за цільовою комплексною програмою наукових досліджень НАН України «Грид-інфраструктура і грид-технології для наукових і науково-прикладних застосувань».

Державний реєстраційний номер роботи 0117U002721

features of ECG; identification; authentication

I. ВСТУП

Як відомо, використання ЕКГ для ідентифікації пацієнта відноситься до біометричних методів ідентифікації. Протягом багатьох років ЕКГ вивчалися і використовувалися в клінічній діагностиці. Але порівняно недавно було запропоновано використовувати ЕКГ в якості біометричної інформації для ідентифікації людей [1-5]. У біометричному підході нам потрібно знайти схожість в ЕКГ - сигналах, що належать одній і тій же людині, і при цьому розпізнавати відмінності в ЕКГ - сигналах, що належать різним людям. Ідентифікація (автентифікація) людини за її електрокардіограмою в 6-ти або 12-ти стандартних відведеннях є актуальною саме зараз, на сучасному етапі розвитку мобільних медичних пристроїв, для ефективної організації телемедичних консультацій.

В роботах [6-8] нами наведені алгоритмічні рішення, які формують системи інформаційних ознак у вигляді керуючих точок параметричного сплайна, які відповідають QRS-комплексу електрокардіограми в трьох ортогональних відведеннях або перших трьох відведеннях. Наступний етап роботи з цією системою ознак має використовувати автоматичний класифікатор, що буде проводити ідентифікацію пацієнта. У даній роботі ми описуємо проведені експерименти для аутентифікації і ідентифікації людини за її ЕКГ,

застосовуючи класифікатор CONFOR, що базується на використанні зростаючих пірамідальних мереж.

II. ПІРАМІДАЛЬНА МЕРЕЖА ТА СИСТЕМА CONFOR

В основу системи CONFOR [9] покладено метод індуктивного формування понять (CONcept FORmation), який базується на використанні зростаючих пірамідальних мереж. Важливою властивістю пірамідальних мереж є їх ієрархічність, що дозволяє природним чином відображати структуру складених об'єктів і родовидові зв'язки "Рис 1". Система CONFOR працює в трьох режимах: навчання, розпізнавання, екзамен. При використанні пірамідальних мереж всі пошукові операції обмежуються відносно невеликим фрагментом мережі, який включає піраміду об'єкт і вершини, безпосередньо з нею пов'язані.

За відсутності точного рішення для об'єкта, який розпізнається, користувачеві надаються всі можливі рішення, в тому числі і рішення «не знаю». Рішення ранжуються відповідно до обчислених ступенів довіри.

Підсистема навчання здійснює виділення закономірностей (знань) які характеризують класи об'єктів [10]. Вхідними даними для підсистеми навчання служить вибірка для навчання, яка включає приклади об'єктів, що характеризуються різними класами. Кожен клас об'єктів має бути представлений деякою кількістю прикладів, достатнім для того, щоб на підставі їх аналізу була виділена закономірність, що характеризує цей клас. Приклади об'єктів задаються ознаковими описами. Ознаковий опис об'єкту повинен включати ім'я, клас, до якого об'єкт належить, і набір значень ознак, які характеризують об'єкт. Вихідні дані для підсистеми навчання – закономірності які характеризують класи об'єктів. Результатом роботи CONFOR'a в режимі навчання є узагальнена модель класів об'єктів, що включає найбільш характерні властивості цих об'єктів.

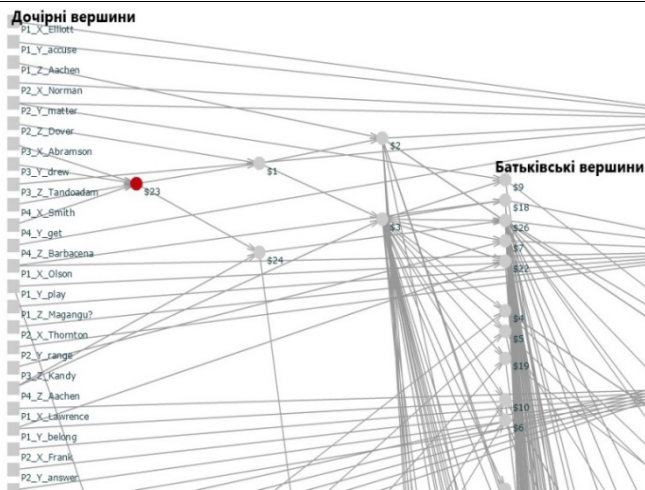


Рис. 1. Побудова мережі для ознак класу, що проходить навчання

Підсистема іспиту призначена для тестування якості навчання комплексу. Вхідними даними для підсистеми іспиту служить екзаменаційна вибірка що включає приклади об'єктів, які не увійшли до вибірки для навчання, але відносно яких відомо, до якого класу вони

належать. Вихідні дані підсистеми - її відповідь відносно класу, до якого віднесена кожен об'єкт, і статистичні дані про кількість правильних, неправильних і невизначених відповідей підсистеми.

Отримані результати зіставляються з інформацією про реальну приналежність об'єктів до відповідних класів, на основі чого можна судити про те, наскільки добре навчений комплекс.

III. ПІДГОТОВКА ДАНИХ

Стандартна електрокардіограма складається з 12 відведень, тобто 12 сигналів. За допомогою спеціальних алгоритмів здійснюється попередня обробка цих сигналів, яка включає в себе: перетворення в три ортогональних відведення, фільтрацію і розмітку - поділ на кардіоцикли і елементи циклів [11]. Тривимірну дугу, що відповідає QRS-комплексу, апроксимуємо параметричними сплайнами. При цьому отримуємо кількість точок, що характеризують форму дуги. Для апроксимації використовуємо канонічний сплайн (cardinal spline), також можна використовувати криву Безьє [6].

Керуючі точки параметричного сплайна і є нашою системою ознак. Це означає, що для проведення експерименту в системі CONFOR необхідно керуючі точки QRS-комплексу, представлені координатами x, y, z, перетворити в символічний вигляд (табл. 1).

ТАБЛИЦЯ I. КЕРУЮЧІ ТОЧКИ QRS-КОМПЛЕКСУ В СИМВОЛЬНОМУ ВИГЛЯДІ

	X	Y	Z
point1	0.0000	0.0000	0.0000
point2	1.0000	1.0000	1.0000
point3	0.3679	0.4030	0.5309
point4	0.1838	0.2105	0.3089

	X	Y	Z
point1	Smith	watch	Dover
point2	Abramson	drew	Baku
point3	Bryan	get	Aachen
point4	Payne	hide	Navapolatsk

Вирішити це питання вдалось, з допомогою спеціальної програми де вихідний файл подавався в додаток, в якому відбувалась робота з англійським словником, який попередньо був підготовлений для виконання задачі. Для осі X словник налічував 1000 англійських слів (імена). Для

осі Y використовували 1000 дієслів і для осі Z був підібраний словник з назвами міст різних країн. Це означає, що кожна координата має свою певну фразу і дозволяє не тільки виконати завдання подачі даних в систему, а є зручним у використанні і запам'ятовуванні слів, що належать до певної точки.

На виході всі координати для CONFOR було представлено в символічному вигляді.

IV. ПРОВЕДЕННЯ ЕКСПЕРИМЕНТІВ

Ми продовжуємо проводити серію експериментів [12-15] на більшій кількості даних. Для проведення наступних експериментів було зареєстровано ЕКГ різних людей, зокрема 18 пацієнтів, функціонально здорових, віком від 25 до 65 років чоловічої і жіночої статі, які знаходилися в різному емоційному та фізичному стані.

Сформувавши систему ознак, для вибірки в підсистему навчання було подано по 100 QRS-комплексів для кожного пацієнта. Для підсистеми іспиту було подано від 65 до 1642 QRS-комплексів, в залежності від кількості отриманих даних певного пацієнта. Потрібно зазначити, що для підсистеми навчання і підсистеми іспиту було подано різні QRS-комплекси.

Перший експеримент був побудований, як «автентифікація пацієнта». Це означає, що в режим підсистеми навчання і в режим підсистеми іспиту для розпізнавання були подані QRS-комплекси одного і того ж пацієнта. Результати автентифікації можна розглянути в таблиці 1. Бачимо, що 50% пацієнтів автентифіковано стовідсотково. Інші 50% пацієнтів з невеликими помилками (табл. 2).

Другий експеримент для 10 пацієнтів побудований, як «ідентифікація пацієнта». В режим навчання підсистеми подавалися 100 QRS-комплексів пацієнта, якого потрібно ідентифікувати.

В режим підсистеми іспиту були подані по 100 QRS-комплексів від кожного з трьох інших пацієнтів і також 100 QRS-комплексів пацієнта, якого потрібно розпізнати з суми всіх даних.

ТАБЛИЦЯ II. РЕЗУЛЬТАТИ АВТЕНТИФІКАЦІЇ ПАЦІЄНТІВ

Клас	Навчальна вибірка (QRS)	Кількість QRS-комплексів в режимі «іспит»	Автентифікація (%)	Неправильно (QRS)	Невизначено (QRS)
Пацієнт 1	100	368	99,73	0	1
Пацієнт 2	100	1455	99,45	0	8
Пацієнт 3	100	468	100	0	0
Пацієнт 4	100	257	91,44	0	22
Пацієнт 5	100	181	100	0	0
Пацієнт 6	30	65	100	0	0
Пацієнт 7	100	380	100	0	0
Пацієнт 8	100	1642	99,88	0	2
Пацієнт 9	100	1391	94,32	0	79
Пацієнт 10	30	92	100	0	0
Пацієнт 11	30	58	100	0	0
Пацієнт 12	100	1218	99,92	0	1
Пацієнт 13	20	77	100	0	0
Пацієнт 14	100	1162	99,57	0	5
Пацієнт 15	100	166	100	0	0
Пацієнт 16	100	1439	99,86	0	2
Пацієнт 17	100	890	98,76	0	11
Пацієнт 18	100	485	100	0	0

Такі процедури проводилися для 10-ти пацієнтів. В результаті ми можемо сказати, що система успішно розпізнала пацієнтів, за мету яких ми ставили розпізнати (табл. 4).

Закономірність ознак може бути представлена у вигляді логічного виразу, який є більш наочним і легко інтерпретується. Успішна ідентифікація пояснюється тим, що повністю виконуються логічні вирази в системі CONFOR, які будуються для кожного пацієнта.

Наприклад, для пацієнта 1 і пацієнта 2 в режимі підсистеми навчання, було отримано наступні кон'юнкції, що характеризують об'єкти даного класу:

Клас: пацієнт 1 – P3_X_Abramson & P3_Y_drew & P4_X_Smith & P4_Y_get.

Клас: пацієнт 2 – P3_X_Abramson & P3_Y_drew & P3_Z_Dover & P4_Y_get & P4_Z_Aachen.

Коли логічні вирази виконуються не повністю, частково, в ході експерименту було виявлено помилку ідентифікації – пацієнт ідентифікований не правильно.

Третій експеримент проводився подібно другому експерименту. В режим підсистеми навчання (табл.3) були подані по 100 QRS-комплексів від 8 пацієнтів і по 30 QRS-комплексів двох інших пацієнтів. В підсистему іспиту подано по 200 інших QRS-комплексів цих пацієнтів. Далі маємо наступні результати.

ТАБЛИЦЯ III. ІНФОРМАЦІЯ МЕРЕЖІ ПІД ЧАС РЕЖИМУ НАВЧАННЯ

Всього вузлів	4487
Об'єктів	860
Класів	10
Головних вузлів	796
Проміжних вузлів	421
Рецепторів	3225
Зв'язків вузлів	7425

Система розпізнала з точністю від 90-100% сім пацієнтів, решта пацієнтів була розпізнана з меншою

ТАБЛИЦЯ ІV. РЕЗУЛЬТАТИ ІДЕНТИФІКАЦІІ ПАЦІЄТІВ

Клас	Навчальна вибірка (QRS)	Кількість QRS – комплексів в режимі «іспит»	Правильно ідентифіковано (QRS)	Неправильно (QRS)	Невизначено (QRS)	Класи підсистеми іспиту
Пацієнт 1	100	368	367	0	1	Пацієнт - 1, Пац- 2, Пац- 3, Пац-4
Пацієнт 2	100	300	300	0	0	Пац- 2, Пац- 1, Пац- 3, Пац-4
Пацієнт 3	100	300	300	0	0	Пац- 3, Пац- 1, Пац- 2, Пац-4
Пацієнт 4	100	250	228	0	22	Пац- 4, Пац- 2, Пац- 1, Пац-5
Пацієнт 5	100	180	180	0	0	Пац- 5, Пац- 4, Пац- 3, Пац-2
Пацієнт 6	30	65	65	0	0	Пац- 6, Пац- 4, Пац- 7, Пац-2
Пацієнт 7	100	360	360	0	0	Пац- 7, Пац- 2, Пац- 4, Пац-5
Пацієнт 8	100	1640	1638	0	2	Пац- 8, Пац- 7, Пац- 5, Пац-4
Пацієнт 9	100	1390	1311	0	79	Пац- 9, Пац- 4, Пац- 5, Пац-7
Пацієнт 10	30	90	90	0	0	Пац- 10, Пац- 4, Пац- 5, Пац-7

точністю, про що свідчать неправильно віднесені QRS-комплекси між пацієнтами. Для вирішення питання правильної ідентифікації великої кількості пацієнтів, необхідно в системі CONFOR підвищити ступінь довіри, коли формуються і виконуються логічні вирази.

ВИСНОВКИ

При використанні пірамідальної мережі з'являється реальна можливість вирішувати практичні дослідницькі завдання, пов'язані з обробкою великих обсягів даних, а саме ідентифікацію людини за її електрокардіограмою.

При наявності відповідної системи ознак, на достатній кількості даних з допомогою системи CONFOR, було проведено серію експериментів – «ідентифікація пацієнтів» і «автентифікація пацієнтів». Отримані результати вказуються на успішне застосування системи, в якості класифікатора. Ідентифікація і автентифікація досягла 100%.

Наразі ми продовжуємо експеримент на масиві електрокардіограм 130-ти пацієнтів, які накопичені у проекті «МедГрид».

ЛІТЕРАТУРА REFERENCES

[1] L.Fainzilberg, T. Potapova, "Computer Analysis and Recognition of Cognitive Phase Space Electrocardiographic Image", Proceeding of 6th International Conference on Computer analysis of Images and Patterns, Prague, 1995, pp 668-673.

[2] L. Biel., O. Petersson, L. Philipson, P. Wide, "ECG Analysis: a New Approach in Human Identification," IEEE Transactions on Instrumentation and Measurement, 2001, vol 3, pp. 808-812.

[3] F. Sufi, I. Khalil, I. Habib, "Polynomial distance measurement for ECG based biometric authentication," Security and Communication Networks, 2010, vol 4, pp. 303-319.

[4] S. Yogendra Narain, "Biometrics Method for Human Identification Using Electrocardiogram," Proceedings of third International Conference, ICB. Alghero, Italy, 2009, pp. 1270-1279.

[5] F. Agrafioti, D. Hatzinakos, "Fusion of ECG sources for human identification," *Control, and Signal Processing: International Symposium on Communications*, 2008, Malta, pp. 1542-1547.

[6] V. Vishnevsky, V. Kalmykov, T. Romanenko, "Approximation of one-, two- and three-dimensional Curves Arcs by Parametric Splines," *Mathematical Machines and Systems*, 2015, No. 4, pp. 160.

[7] V. Vishnevsky, T. Romanenko, L. Kizub, "Experimental Verification of Possibility of Human Identification by the Electrocardiogram," *Proceedings of 5TH International Conference on Application of Information and Communication Technology and Statistics in Economy and Education*, Sofia, Bulgaria, 2015, pp. 318.

[8] V. Vishnevsky, "Grid System for Mass Storage and Processing of Digital Electrocardiograms," *Ukrainian Journal of Telemedicine and Medical Telematics*, 2013, vol. 11, No. 1, pp. 202-208.

[9] The Association of Developers and Users of Intelligent Systems [Online]. Available: <http://www.aduis.com.ua>.

[10] V. Gladun, V. Velichko, "Instrumental Complex of Decision-Making Support on the Basis of the Network Model of the Subject Object Field," *Decision Support Systems. Theory and practice*, Kiev, 2012, pp. 128.

[11] I. Tchaikovsky, "Analysis of the electrocardiogram in one, six and twelve leads with point of view of information value: electrocardiographic cascade," *Clinical Informatics and Telemedicine*, 2012, vol 2, pp. 102-106.

[12] V. Vishnevsky, T. Romanenko, L. Kizub, "Biometric Identification Using Electrocardiograms," in *Information Technologies and Computer Modelling: International Scientific and Practical Conference ITCM-2015*, Ivano-Frankivsk- Vinnytsia, 2015, pp. 130-131.

[13] V. Vishnevsky, T. Romanenko, L. Kizub, "Identifying Individuals Using Electrocardiograms and their characteristics," in *Visnyk of Vinnytsia Polytechnic Institute*, 2016, vol 5, pp. 7-10.

[14] V. Vishnevsky, V. Velychko, L. Kizub, "Applications of Growing Pyramidal Network for Solving to the Problem of Human Identification with Using Electrocardiogram," in *Information Technologies and Computer Modelling: International Scientific and Practical Conference ITCM-2017*, Ivano-Frankivsk, 2017, pp. 104-107.

[15] V. Velychko, L. Kizub, "Using the Pyramidal Network in the Model Experiment of Human Identification by its ECG," in *Mathematical and Simulation Modeling Systems: International Scientific and Practical Conference - 2017*, Chernigov, 2017, pp. 79-82.